DOCUMENT RESUME

ED 346 118 TM 018 365

AUTHOR Reshetar, Rosemary A.; Swaminathan, Hariharan
TITLE A Comparison of the Log-Linear and Weighted Least Squares Approaches for the Analysis of Categorical Data.
PUB DATE Apr 92
NOTE 23p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 20-24, 1992).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Chi Square; *Classification; *Comparative Analysis; Effect Size; *Hypothesis Testing; *Least Squares Statistics; Mathematical Models; Simulation; *Statistical Distributions
IDENTIFIERS *Categorical Data; Contingency Tables; *Log Linear Models; Power (Statistics); Type I Errors; Weighted Data

ABSTRACT

This study compared the model of J. E. Grizzle, C. F. Starmer, and G. G. Koch (GSK, 1969) and log-linear model-based approaches for testing hypotheses in r x c contingency tables. Tables were simulated under various conditions of table, sample, row-effect size, and column-effect size. Test statistics for column (main) and interaction effects were calculated using: (1) the GSK linear model, untransformed proportion; (2) GSK linear model, logarithm of the proportion; (3) GSK linear model, log-odds log (p/1-p) ; and (4) log-linear model. Type I error rates, distribution of the test statistics, and relative power of the procedures were studied. The log-linear method provided more accurate distributions of the chi square test statistics than did the GSK methods and yielded Type I error rates closer to the expected 5% and 1% levels. The GSK methods using linear and logarithmic response functions yielded less precise distributions of chi square test statistics and slightly inflated Type I error rates. This trend was most noticeable with the linear response function of untransformed proportions. Based on these findings, if the GSK method is used, a logarithmic transformation of the response function is recommended. It is also recommended that a more stringent Type I error rate be used with GSK methods since observed Type I error rates were higher than expected. Nine tables and six graphs present analysis results, and there is a five-item list of references. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED346118

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

ROSEMARY RESHETAR

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

A Comparison of the Log-Linear and Weighted Least Squares Approaches For the Analysis of Categorical Data

Rosemary A. Reshetar
American Board of Internal Medicine

Hariharan Swaminathan
University of Massachusetts

Paper presented at the Annual Meeting of the American Educational Research Association, 1992, San Francisco, CA.

TM018365

BEST COPY AVAILABLE

A Comparison of the Log-Linear and Weighted Least Squares Approaches for the Analysis of Categorical Data

Rosemary Reshetar, American Board of Internal Medicine
Hariharan Swaminathan, University of Massachusetts

Abstract

This study compared the Grizzle, Starmer & Koch (GSK) and log-linear model-based approaches for testing hypotheses in *r x c* contingency tables. Tables were simulated under various conditions of table, sample, row-, and column-effect sizes. Test statistics for column (main) and interaction effects were calculated using: (i) GSK linear model, untransformed proportion; (ii) GSK linear model, logarithm of the proportion; (iii) GSK linear model, log-odds [log (p/1-p)]; and (iv) log-linear model. Type I error rates, distribution of the test statistics, and relative power of the procedures were studied.

The log-linear method provided more accurate distributions of χ^2 test statistics than did the GSK methods, and yielded Type I error rates closer to the expected 5% and 1% levels. The GSK methods using linear and logarithmic response functions yielded less precise distributions of χ^2 test statisticss and slightly inflated Type I error rates. This trend was most noticeable with the linear response function of untransformed proportions.

Based on findings of this study, if the GSK method is utilized, a logarithmic transformation of the response function is recommended. It is also recommended that a more stringent Type I error rate be used with GSK methods since observed Type I error rates were higher than expected.

A Comparison of the Log-Linear and Weighted Least Squares Approaches for the Analysis of Categorical Data

Rosemary Reshetar, American Board of Internal Medicine
Hariharan Swaminathan, University of Massachusetts

Multiway contingency tables may arise in experimental work when data on two or more variables are assigned to nominal categories. Data of this type occur, for example, when the members of a sample are cross-classified into categories representing different characteristics. The researcher may be interested in studying the association among the attributes defining the ways of classification as well as other aspects of the structure of the data.

The most common approach for analyzing such data uses the Pearson chi-square statistic for testing the hypothesis of independence. The Pearson chi-square test is a general test of independence and is thus appropriate only when the research question is one of association between the categories of interest. When the researcher is interested in testing structural hypotheses pertaining to row and/or column effects analogous to main effects hypotheses in analysis of variance, a model-based approach is necessary.

Background

Using a model-based approach, once a table is constructed based on assignment to categories, hypotheses about the relationship between two or more variables may be tested by analysis of the probability distribution of the observed frequencies or of some transformation of these frequencies. Two model-based approaches for the testing of structural hypotheses are the log-linear model approach (see, for example, Goodman, 1978) and the weighted least square (WLS) method suggested by Grizzle, Starmer, and Koch (GSK) (1969). The log-linear approach utilizes maximum likelihood (ML) iterative procedures for estimating parameters. The GSK linear model approach is non-iterative and uses the weighted least squares procedure for estimating parameters. The two methods are asymptotically equivalent.

With the GSK method a hypothesized model of the form $\mathbf{F} = \mathbf{X}\boldsymbol{\beta}$ where $\mathbf{X}$ is a known design matrix can be tested. The WLS estimator of $\boldsymbol{\beta}$ is $\mathbf{b} = (\mathbf{X}'\mathbf{S}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{S}^{-1}\mathbf{F}$, where $\mathbf{F}$ is a function of the observed cell frequencies and $\mathbf{S}$ is the estimated sample variance-covariance matrix of $\mathbf{F}$. The fit of the model may be tested using the test statistic $\chi^2 = (\mathbf{F} - \mathbf{Xb})'\mathbf{S}^{-1}(\mathbf{F} - \mathbf{Xb})$. If the model fits, linear hypotheses about the $\boldsymbol{\beta}$ can be tested as H_o: $\mathbf{C}\boldsymbol{\beta} = \mathbf{0}$, where $\mathbf{C}$ is a full rank matrix of arbitrary constants. The test statistic of the hypothesis is given by $\chi^2 = (\mathbf{Cb})'[\mathbf{C}(\mathbf{X}'\mathbf{S}^{-1}\mathbf{S})^{-1}\mathbf{C}']^{-1}\mathbf{Cb}$. In developing a structural model, a function of the cell probabilities $\mathbf{F} = \mathbf{F}(\pi_{ij})$ is developed and used to test hypotheses. The functions that can be formed involve linear, logarithmic, exponential, or combined transformations of the cell probabilities (see Forthofer & Lehnen, 1981).

Purpose of the Study

Little is known about the relative merits of the log-linear and GSK procedures. Drew (1985) examined the performance of chi-square tests associated with the WLS approach for 2xc tables and small sample sizes. He found that tests using logits or complementary log-logs were conservative. As expected, it was found that the accuracy of the chi-square approximation increased as sample

size increased. Drew's investigation was limited to 2 x c tables (the binomial case). Since the situation may be different in the case of general r x c tables, it is important to study the relative merits of the two procedures in the general case.

The purpose of this study is to compare the GSK linear model approach using WLS procedures with the log-linear model approach for testing hypotheses of interest in r x c contingency tables. The following four estimation methods and response function pairs were studied: (i) GSK linear model, untransformed proportion (p); (ii) GSK linear model, logarithm of the proportion [log(p)]; (iii) GSK linear model, log-odds [log p/(1-p)]; and (iv) log-linear model (using ML estimation).

Design of the Study

In order to compare the procedures, a simulation study was carried out. The factors manipulated were sample size, table size, and row and column effect sizes. Two-way (rxc) tables were simulated so that there was either a column effect or both column and row effects. Tables with only row effects were not simulated because these are identical to tables with only column effects. Three sample sizes (N=250, 500, and 1000) and three table sizes (4x2, 4x3, and 4x4) were studied. For all tables, five levels of column effect sizes (CES) were produced, with effect (1) corresponding to no column effect. The column marginal probability parameters used for the simulation are given in Table 1.

Table 1 Column Marginal Probability Parameters for Each Level of Column Effect Size
$\pi_{.j}$ parameters, where $\pi_{.j}$ = column marginal probability parameter for column j

Column Effect Size	2 Column Tables		3 Column Tables			4 Column Tables			
	$\pi_{.1}$	$\pi_{.2}$	$\pi_{.1}$	$\pi_{.2}$	$\pi_{.3}$	$\pi_{.1}$	$\pi_{.2}$	$\pi_{.3}$	$\pi_{.4}$
1)*	.5000	.5000	.3333	.3333	.3333	.2500	.2500	.2500	.2500
2)	.4850	.5150	.3183	.3333	.3483	.2350	.2450	.2550	.2650
3)	.4700	.5300	.3033	.3333	.3633	.2200	.2400	.2600	.2800
4)	.4550	.5450	.2883	.3333	.3783	.2050	.2350	.2650	.2950
5)	.4400	.5600	.2733	.3333	.3933	.1900	.2300	.2700	.3100

*Note: Size 1=No column effect

For each one of the column effect sizes, row effects were specified in three ways: (1) no row effect, (2) some row effect, and (3) a larger row effect. The row marginal probability parameters are given in Table 2. Individual cell probability parameters used for the simulation were created by multiplying the column and row marginal parameters.

Table 2 Row Marginal Probability Parameters for Each Level of Row Effect Size
$\pi_{i.}$ = row marginal probability parameter for row i

Row Effect Size	$\pi_{1.}$	$\pi_{2.}$	$\pi_{3.}$	$\pi_{4.}$
1)	.25	.25	.25	.25
2)	.22	.24	.26	.28
3)	.19	.23	.27	.31

All factors were completely crossed (three sample sizes by three table sizes by three row effect sizes by five column effect sizes) yielding a total of 135 tables. The properties of the test statistics were studied by carrying out one thousand replications of each table.

Using each of the four estimation method/response function pairs given above, chi-square test statistics for tests of goodness of fit and no column main effects were computed. For example, for the 4x4 tables we may first test the goodness of fit of a model with only column and row effect parameters given as $\mathbf{F}(\pi) = \mathbf{X}\,\beta$, where the elements of $\mathbf{X}$ indicate the factor level combination being considered. The parameters of interest in this case are

β_0 = mean term
β_1 = differential column 1 effect
β_2 = differential column 2 effect
β_3 = differential column 3 effect
β_4 = differential row 1 effect
β_5 = differential row 2 effect
β_6 = differential row 3 effect

Since only row and column effects were simulated, the test of interaction also serves as the test of the goodness of fit of the model.

To test column effects, the β parameters corresponding to column effects (for 2-column tables, β_1; for 3-column tables, β_1 and β_2; and for 4-column tables, β_1, β_2 and β_3) were set equal to zero. This was done by testing the hypothesis $\mathbf{H_o}$: $\mathbf{C}\beta = \mathbf{0}$, where the constraint matrix $\mathbf{C}$ for the 4-column tables is

$$\mathbf{C} = \begin{matrix} 0\,1\,0\,0\,0\,0\,0 \\ 0\,0\,1\,0\,0\,0\,0 \\ 0\,0\,0\,1\,0\,0\,0 \end{matrix}.$$

Results

Type I Error Rates

Tests of column effects. For each table simulated, values for χ^2 tests of no column effect were computed using each of the four estimation methods. The percentage of χ^2 values significant at $\alpha = .05$ and $\alpha = .01$ were then tabulated. Table 3 gives summaries of the Type I error rates for the four methods.

Table 3 Summary of Type I Error Rates
χ^2 Tests of Column Effects, Tables with No Column Effects (CES = 1)

	.05 Level		.01 Level	
Method	Range of Error	Mean Error	Range of Error	Mean Error
GSK Linear	5.7% to 8.9%	7.21%	1.2% to 2.6%	1.77%
GSK Log(p)	4.6% to 7.5%	6.06%	0.6% to 2.1%	1.25%
GSK Log-Odds	4.5% to 7.4%	6.10%	0.6% to 2.2%	1.26%
ML Log-Linear	3.6% to 6.1%	5.08%	0.3% to 1.8%	1.00%

The results were similar for tables simulated with no row effects and tables simulated with row effects; and detailed results for all CES are presented for tables simulated with row effect size (3) in Table 4 for $\alpha = .05$ and Table 5 for $\alpha = .01$ (see next pages). With the GSK linear model, the rates exceeded the expected 5% (1%) level in most cases, and the mean rate of 7.21% (1.77%) was much higher than expected. Very similar results were obtained for the GSK log and the GSK log-odds methods, with mean Type I error rates of 6.06% (1.25%) and 6.10% (1.26%). In contrast the log-linear model provided lower Type I error rates with a mean value of 5.08% (1.00%) very close to the expected value.

No systematic variations due to row effect size, sample size or table size were noted for any of these methods. In general, a trend was detected such that the patterns of change in the Type I error rates were similar between all methods, where increases and decreases in one method corresponded to increases and decreases in the other methods. Also each of the GSK methods were typically biased in that the observed Type I error rates were greater than the expected 5% and 1%. This bias was greater when the linear response function was used than when either logarithmic response function was used. In most cases, the GSK methods gave Type I error rates greater than the ML log-linear method which overall proved to be much closer to the expected 5% and 1%.

Table 4

Percentage of X^2 Column Values Significant at .05 Level
For Tables Simulated With Row Effect Size (3)

		Size: 4x2			Size: 4x3			Size: 4x4		
Column Effect Size	Sample Size: / Method	250	500	1000	250	500	1000	250	500	1000
1	GSK Linear	7.1	7.1	6.7	7.3	7.3	6.5	8.9	7.9	7.5
	GSK Log(p)	5.9	6.5	6.9	5.7	6.4	5.9	5.2	6.6	6.8
	GSK Log-odds	6.3	6.7	6.5	5.7	6.4	5.9	5.1	6.7	6.9
	ML Log-linear	4.9	5.8	4.7	4.9	5.0	5.2	4.9	6.0	5.3
2	GSK Linear	11.6	12.7	19.2	9.1	13.6	18.4	13.7	13.4	22.6
	GSK Log(p)	10.8	12.1	19.7	7.1	12.3	17.6	10.2	11.8	20.5
	GSK Log-odds	11.4	12.4	19.7	7.1	12.5	17.8	10.1	12.1	20.5
	ML Log-linear	8.9	10.9	15.4	6.0	10.0	15.2	9.4	10.7	18.6
3	GSK Linear	19.0	31.2	52.3	21.3	34.1	57.3	24.8	41.9	69.6
	GSK Log(p)	18.4	31.9	52.4	18.8	32.2	58.2	19.7	38.2	69.9
	GSK Log-odds	18.6	32.4	52.4	18.9	32.3	58.3	19.7	38.5	70.4
	ML Log-linear	15.7	29.6	47.5	17.5	30.6	54.1	18.8	37.1	67.2
4	GSK Linear	33.8	53.6	82.5	34.9	62.5	88.3	43.7	73.4	97.1
	GSK Log(p)	33.7	54.2	82.1	29.8	62.3	89.0	37.0	74.3	97.0
	GSK Log-odds	34.5	54.7	82.3	30.3	62.3	88.7	36.8	73.9	97.1
	ML Log-Linear	30.3	50.5	78.7	28.8	59.8	87.5	36.8	73.6	96.6
5	GSK Linear	50.2	80.2	98.2	58.2	85.3	99.0	71.8	93.9	100.0
	GSK Log(p)	51.3	80.1	98.2	57.8	86.5	99.2	67.1	94.2	100.0
	GSK Log-odds	51.7	80.9	98.3	58.3	86.6	99.3	67.2	94.4	100.0
	ML Log-Linear	46.4	78.6	97.5	56.1	84.6	99.1	67.2	93.2	100.0

Table 5

Percentage of X^2 Column Values Significant at .01 Level
For Tables Simulated With Row Effect Size (3)

		Size: 4x2			Size: 4x3			Size: 4x4		
Column Effect Size	Sample Size: / Method	250	500	1000	250	500	1000	250	500	1000
1	GSK Linear	1.6	1.3	1.7	1.4	1.6	1.6	2.4	1.8	2.3
	GSK Log(p)	1.6	1.6	1.3	.6	1.2	.9	.7	1.0	1.5
	GSK Log-odds	1.6	1.6	1.3	.6	1.2	.9	.7	1.0	1.5
	ML Log-linear	1.1	.9	1.0	.3	.6	.7	.9	.8	1.2
2	GSK Linear	3.1	3.9	7.9	2.6	4.3	6.8	5.4	4.9	10.0
	GSK Log(p)	2.6	3.8	7.6	2.3	3.9	5.8	2.1	4.3	8.3
	GSK Log-odds	2.7	3.6	7.5	2.3	3.9	6.0	2.2	4.3	8.0
	ML Log-linear	2.0	2.6	5.4	2.2	2.9	4.8	2.0	3.5	7.5
3	GSK Linear	7.9	14.3	31.6	9.3	16.5	34.7	10.2	21.3	47.4
	GSK Log(p)	6.4	14.1	32.8	7.1	14.6	34.9	7.5	18.0	45.4
	GSK Log-odds	6.7	14.1	32.1	7.4	14.5	35.1	7.9	18.1	45.5
	ML Log-linear	5.8	11.2	27.2	6.6	12.4	31.4	6.9	16.6	42.0
4	GSK Linear	14.0	31.2	63.3	14.4	40.1	74.1	23.4	54.5	89.5
	GSK Log(p)	13.9	30.8	63.9	11.6	38.8	75.6	17.4	52.3	89.5
	GSK Log-odds	13.9	30.9	63.5	12.0	38.9	75.5	17.2	52.2	89.5
	ML Log-Linear	12.1	26.9	58.4	10.2	34.4	71.5	16.9	49.2	88.2
5	GSK Linear	28.6	60.5	91.4	37.1	71.0	97.0	49.1	83.9	99.7
	GSK Log(p)	26.6	60.5	92.4	31.9	69.5	96.4	42.3	83.0	99.7
	GSK Log-odds	27.0	60.9	92.2	32.3	69.7	96.4	42.5	83.2	99.7
	ML Log-Linear	23.7	56.8	90.1	30.7	66.0	95.9	40.5	81.8	99.7

Tests of goodness of fit. For each of the tables, χ^2 tests of goodness of fit or no interaction were conducted. Tables were simulated with no multiplicative interaction. Ranges of Type I error rates for these tests are given in Table 6, and detailed results for tables simulated with row effect size (3) are presented in Table 7 for $\alpha = .05$ and Table 8 for $\alpha = .01$.

Table 6 Summary of Type I Error Rates
χ^2 Tests of Interaction Effects, All Tables

	.05 Level		.01 Level	
Method	Range of Error	Mean Error	Range of Error	Mean Error
GSK Linear	5.7% to 11.9%	7.97%	1.1% to 3.9%	2.03%
GSK Log(p)	4.5% to 9.2%	6.62%	0.6% to 2.6%	1.46%
GSK Log-Odds	4.1% to 9.5%	6.75%	0.6% to 2.7%	1.50%
ML Log-Linear	3.6% to 6.6%	5.21%	0.4% to 2.0%	1.09%

With all of the GSK methods, the Type I error rates were often higher than expected; this was most noticeable with the GSK linear model. In contrast, the log-linear model yielded rates much closer to the expected values of 5% and 1%. No systematic variations due to sample size or table size were noted. For the tables simulated with larger row and column effects (see Tables 7 and 8, CES=5), the GSK linear model yielded noticeably high Type I error rates. It should be pointed out that although the tables were generated without multiplicative interaction, a large row and column effect could result in additive interaction. If additive interaction is present, the difference between column probabilities is different across rows. For example, with the 4x2 tables generated with row effect size (3) and column effect size (5) the differences between the probabilites for columns 1 and 2 are .0228 for the first row and .0372 for the fourth row. The GSK linear model appears to be sensitive to additive interaction and hence yields high Type I error rates. For further discussion of additive and multiplicative interaction, the reader is referred to Forthofer & Lehnen (1981). In general, results of comparisons of Type I error rates were similar for both tests of no column main effect and no interaction.

Null Distributions of the Test Statistic

To determine if the null distribution of the test statistic was indeed χ^2, the distribution of the test statistic was compared to the theoretical χ^2 distribution.

Distribution of test statistics for column effects. For χ^2 tests of no column effects, tests of distribution fit were done for tables where the column effect size equaled one. With the 4x2 tables, none of the methods provided distributions that fit the theoretical distribution. This may be due to imprecision of test statistic estimates when values are very small. With the 4x3 and 4x4 tables, in most of the cases only the three GSK methods provided distributions comparable to the theoretical distributions. With the log-linear model, although the Type I error rates at the .05 and .01 levels agreed with expected values, the distribution did not have the χ^2 distribution.

Table 7

Percentage of X^2 Interaction Values Significant at .05 Level
For Tables Simulated With Row Effect Size (3)

		Size: 4x2			Size: 4x3			Size: 4x4		
Column Effect Size	Sample Size: / Method	250	500	1000	250	500	1000	250	500	1000
1	GSK Linear	9.1	8.1	6.9	5.9	7.3	6.9	7.9	7.6	9.0
	GSK Log(p)	7.9	7.8	6.7	4.5	7.7	6.6	5.3	6.5	7.7
	GSK Log-odds	7.9	7.9	6.8	4.7	7.7	6.5	5.3	6.5	7.9
	ML Log-linear	5.9	5.3	4.1	3.6	5.5	5.3	5.7	5.5	6.1
2	GSK Linear	7.4	6.5	8.8	6.1	7.1	8.0	8.9	7.8	7.6
	GSK Log(p)	6.2	6.2	7.7	4.5	6.3	7.6	6.0	6.2	7.0
	GSK Log-odds	6.6	6.2	7.9	4.8	6.2	7.6	6.2	6.2	7.1
	ML Log-linear	4.3	5.2	5.7	3.7	5.0	5.9	6.5	5.7	5.3
3	GSK Linear	6.8	9.2	8.9	8.3	8.5	10.4	8.5	7.7	8.7
	GSK Log(p)	6.3	8.0	7.4	6.3	6.6	7.9	4.9	5.5	6.6
	GSK Log-odds	6.3	8.2	7.2	6.5	7.0	7.9	4.9	5.6	6.6
	ML Log-linear	4.7	5.2	4.9	5.1	5.5	5.2	5.7	5.0	5.3
4	GSK Linear	8.6	9.3	8.1	8.1	9.2	8.5	9.3	8.9	9.6
	GSK Log(p)	6.7	8.3	6.6	6.4	6.8	7.0	5.7	6.6	7.1
	GSK Log-odds	6.9	8.1	6.5	6.5	7.0	6.8	6.3	6.7	7.1
	ML Log-Linear	4.8	5.3	4.5	5.8	4.9	5.0	6.3	5.4	5.7
5	GSK Linear	10.5	9.5	11.9	7.7	9.3	9.7	8.6	9.0	10.3
	GSK Log(p)	9.0	6.9	7.1	4.9	6.6	6.1	4.1	5.7	4.8
	GSK Log-odds	9.5	7.0	7.0	5.3	6.6	6.1	4.3	5.8	5.4
	ML Log-Linear	6.3	4.7	5.3	4.8	5.3	4.4	5.2	5.3	3.8

Table 8

Percentage of X^2 Interaction Values Significant at .01 Level
For Tables Simulated With Row Effect Size (3)

Column Effect Size	Method \ Sample Size:	Size: 4x2 250	Size: 4x2 500	Size: 4x2 1000	Size: 4x3 250	Size: 4x3 500	Size: 4x3 1000	Size: 4x4 250	Size: 4x4 500	Size: 4x4 1000
1	GSK Linear	2.3	2.1	1.4	1.2	1.9	1.9	1.9	1.7	2.2
	GSK Log(p)	1.8	2.3	1.4	.7	1.7	1.7	.8	1.5	1.9
	GSK Log-odds	1.9	2.3	1.4	.7	1.7	1.7	.8	1.5	2.0
	ML Log-linear	1.2	1.1	.7	.5	1.3	1.1	1.2	1.3	1.5
2	GSK Linear	2.1	2.4	2.3	1.5	1.9	2.1	2.4	1.9	1.6
	GSK Log(p)	1.6	2.1	2.0	1.1	1.4	1.7	1.7	.9	1.2
	GSK Log-odds	1.9	2.4	2.2	1.2	1.6	1.7	1.8	.9	1.3
	ML Log-linear	.7	1.3	1.3	1.0	1.4	1.3	1.8	.7	1.0
3	GSK Linear	1.2	2.3	2.2	2.0	2.2	1.5	2.3	1.4	2.2
	GSK Log(p)	1.0	1.6	1.8	1.2	2.2	1.1	1.1	.8	1.3
	GSK Log-odds	1.0	1.6	1.6	1.4	2.2	1.1	1.2	.8	1.3
	ML Log-linear	.7	.8	1.0	1.4	1.5	.7	1.3	.7	1.1
4	GSK Linear	2.1	2.4	2.1	2.6	1.7	2.6	2.0	2.6	2.7
	GSK Log(p)	1.2	1.9	1.3	1.6	1.0	1.4	.6	1.2	1.6
	GSK Log-odds	1.4	1.9	1.4	1.6	1.2	1.4	.9	1.2	1.7
	ML Log-Linear	.8	1.5	.7	1.5	.7	.7	1.1	1.2	1.2
5	GSK Linear	2.6	2.9	3.9	2.3	2.8	2.8	2.4	2.4	2.8
	GSK Log(p)	1.7	1.5	2.2	1.1	1.6	1.5	.7	1.1	1.2
	GSK Log-odds	2.1	1.7	2.4	1.2	1.7	1.6	.8	1.2	1.2
	ML Log-Linear	.9	1.1	1.0	1.1	1.1	1.1	1.3	1.0	.7

Distribution of test statistics for interaction effects. For tests of interaction, comparisons of observed and theoretical χ^2 distributions were done for all tables with column effect size one. The probability of the χ^2 test for comparison of the observed and theoretical distributions are given in Table 9. For the 4x3 and 4x4 tables, the GSK logistic methods and the log-linear method generally provided acceptable distributions, with the log-linear method most often providing an acceptable distribution. The GSK linear method generally did not provide acceptable distributions. With the 4x2 tables, only the log-linear method provided distributions comparable to the theoretical distributions.

Power

Strictly speaking, given that Type I error rates for the GSK methods are higher than expected, discussions of power are not appropriate. However it is still useful to examine performance of the methods when the null hypothesis is false. When the column effect size is not equal to 1, the null hypothesis of no column effect is false; and the larger the column effect size (CES), the more false the null hypothesis. Similar results were found for tables simulated with no row effects and tables simulated with row effects; thus Tables 4 and 5 (see pages 5 and 6) are used as a representative example. The percentage of χ^2 column values significant at the .05 or .01 level, within each column effect size, is given in Tables 4, and 5. Figures 1 and 2 (pages 12 and 13) graphically depict these results for the 4x4 tables. In these situations where the null hypothesis is false, two factors are influential in the percentage rejection rates. As expected, when the sample size increases and the CES increases the percentage of significant values increases.

A perusal of these results reveals that the three GSK methods generally yielded rejection rates greater than the ML log-linear method did. This is most evident for tables where the rates of rejection ranged from 20% to 80% and can be seen clearly with medium effect sizes (e.g., CES ≥ 2 and CES ≤ 4). In general, the rates of rejection were comparable between methods.

Discussion and Recommendations

For tests of column effects, the test statistics based on the log-linear method yielded Type I error rates closer to the expected 5% and 1% levels in comparison to the GSK methods. The linear response function of the proportions yielded the poorest results; i.e. had the highest Type I error rate.

Similar trends were found for tests of interaction. The GSK procedures yielded higher Type I error rates that the log-linear procedure, with the linear response function yielding the highest Type I error rates. The log-linear method yielded distributions that were closer to the theoretical distribution in comparison to the GSK methods. One reason why the linear function yields highly inflated Type I error rates is that while the data was simulated to have no multiplicative interaction, additive interactions were present. In general, for testing hypotheses of no multiplicative interaction, the more familiar Pearson χ^2 procedure is recommended.

One of the desirable aspects of the GSK method is its flexibility and ease for creating tests of structural hypotheses. The GSK model-based approach may be used with any linear, logarithmic or exponential transformation of the response function. Certain transformations may be advantageous when questions about ordered response data arise. With the GSK method, one of the categorical variables can be modeled as a dependent variable, lending more flexibility to the types of hypotheses that can be tested.

Table 9

Probabilities Resulting From X^2 Tests Comparing
Observed Distributions to Theoretical Distributions

Row Effect Size	Sample Size: / Method	Size: 4x2 250	Size: 4x2 500	Size: 4x2 1000	Size: 4x3 250	Size: 4x3 500	Size: 4x3 1000	Size: 4x4 250	Size: 4x4 500	Size: 4x4 1000
1	GSK Linear	.001	.012	.091	<.001	.087	.117	.009	<.001	.036
	GSK Log(p)	.031	.027	.038	.351	.759	.409	.162	.572	.130
	GSK Log-odds	.019	.105	.336	.555	.366	.778	.265	.302	.223
	ML Log-linear	.030	.782	.107	.848	.903	.179	.555	.244	.845
2	GSK Linear	.100	<.001	<.001	.083	<.001	.006	.053	.248	.027
	GSK Log(p)	.165	.002	<.001	.311	.008	.175	.050	.034	.109
	GSK Log-odds	.087	.001	<.001	.297	.001	.119	.032	.162	.377
	ML Log-linear	.848	.892	.260	.773	.403	.453	.414	.085	.521
3	GSK Linear	.012	.013	<.001	.064	<.001	.510	.078	.036	.001
	GSK Log(p)	.055	.042	<.001	.156	.475	.561	.001	.005	.003
	GSK Log-odds	<.001	.147	.011	.204	.809	.678	.219	.053	.013
	ML Log-linear	.941	.898	.117	.361	.673	.963	.044	.876	.538
	ML Log-Linear									

18

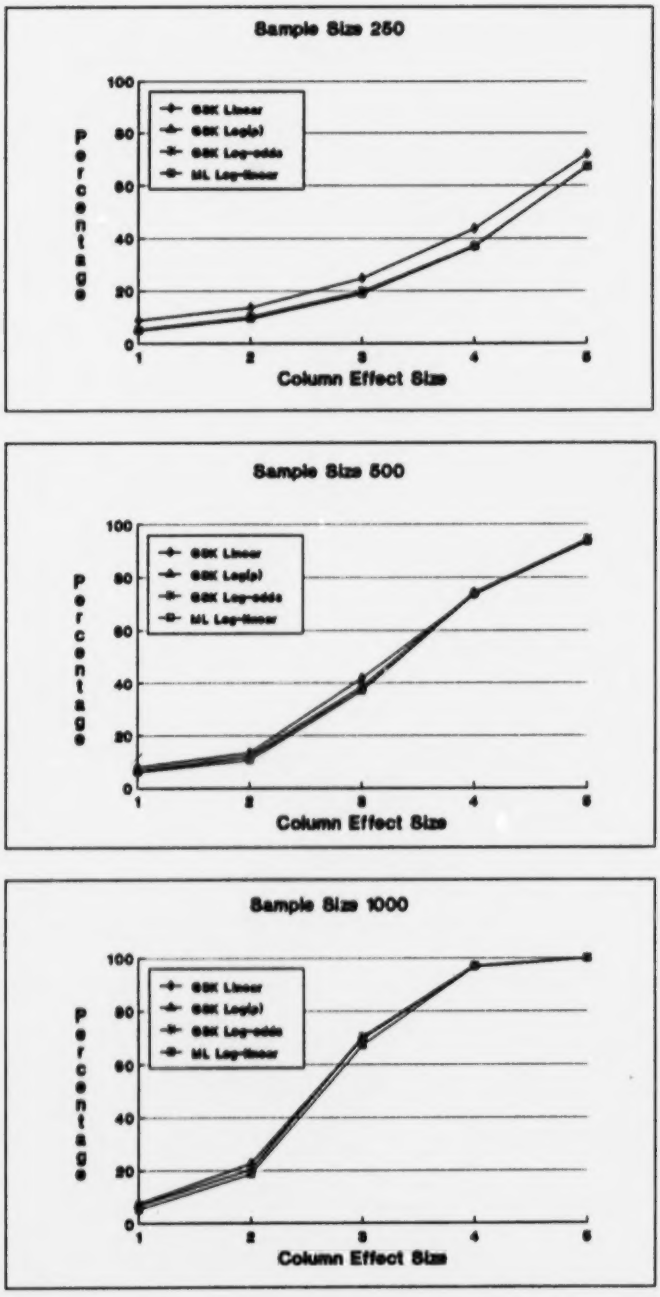

Figure 1
Percentage of Sign. Chi-Square Column Values at .05 Level
4x4 Tables Simulated with Row Effect Size (3)

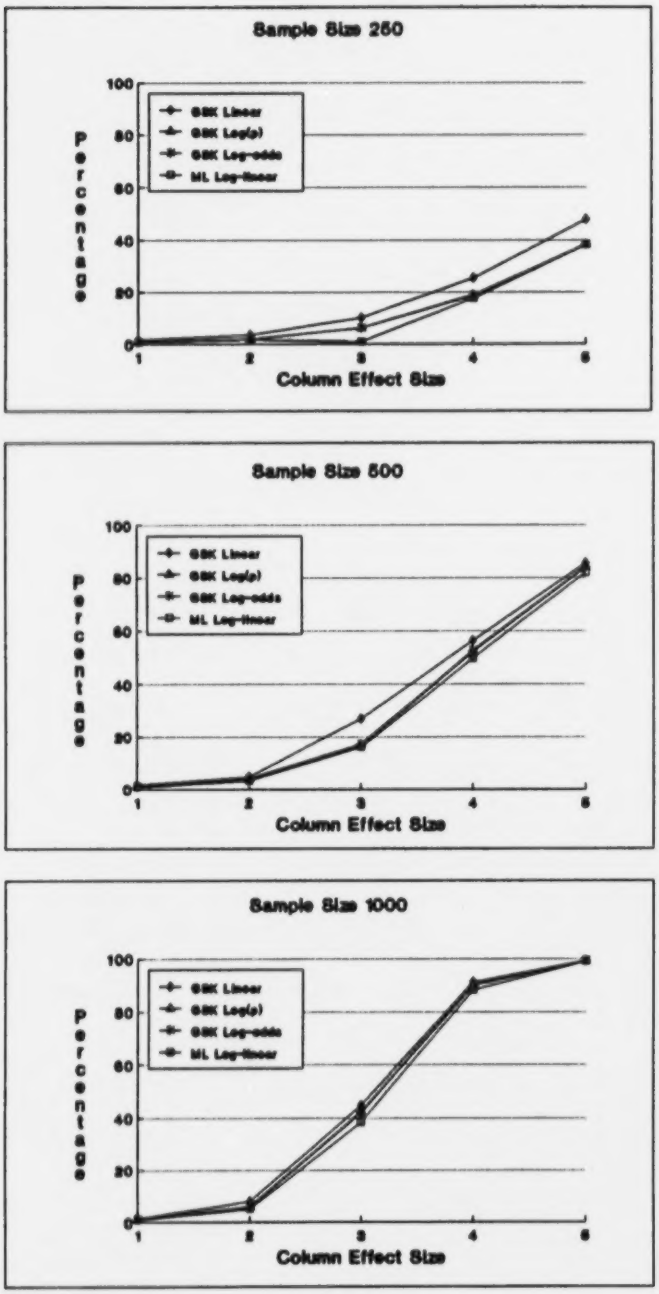

Figure 2
Percentage of Sign. Chi-Square Column Values at .01 Level
4x4 Tables Simulated with Row Effect Size (3)

Log-linear models may also be applied to tests of specific hypotheses using a model-based approach. Discussion of procedures for nonstandard log-linear models is given by Rindskopf (1990) where log-linear models analogous to ANOVA models are presented as special cases of the generalized linear model. Within this context the flexibility of model building discussed for WLS procedures is available. Flexibility does not exist for changing the response function; the logarithm of the cell frequencies is the dependent variable.

With either GSK or log-linear methods the selection of a model and response function that tests the stated hypothesis of interest is important. Based on findings of this study, if the GSK method is utilized, a logarithmic transformation of the response function is recommended. It is also recommended that a more stringent Type I error rate be used with GSK methods since observed Type I error rates were higher than expected.

References

Drew, J. H. (1985). Performance of the weighted least squares approach to categorical data analysis. Comm. Statistics--Theory and Methodology, 14(8), 1963-1979.

Forthofer, R. N., & Lehnen, R. G. (1981). Public program analysis: A new categorical data approach. Belmont, CA: Lifetime Learning Publications.

Goodman, L.A. (1978). Analyzing qualitative/categorical data: Log-linear models and latent-structure analysis. Ed. by Jay Magidson, Cambridge, MA: Abt Books.

Grizzle, J. E., Starmer, C. F., & Koch, G. G. (1969). Analysis of categorical data by linear models. Biometrics, 25, 489-504.

Rindskopf, D. (1990). Nonstandard log-linear models. Psychological Bulletin, 108(1), 150-162.